General model of railway transportation capacity

H. E. Boysen
Department of Transport Science,
Royal Institute of Technology (KTH), Sweden

Abstract

A general model of railway transportation capacity has been developed, which models mass transportation capacity and volume transportation capacity per unit time. It can also be adapted to passenger transportation. The model is easy to use and can handle trade-offs between conflicting effects. It is a suitable tool in capacity planning for preliminary assessment of the capacity effects of candidate investment and operating scenarios, and can point to areas where detailed analysis with other methods should be applied. Application of the capacity model to cases of higher axle loads and higher speeds in freight transportation showed significant and mixed capacity effects that would not have been obvious otherwise.
Keywords: capacity, efficiency, fluidity, freight, model, rail, railroad, railway, train.

1   Introduction

Many of the world's railways face growing demands for both freight and passenger transportation.

Recent statistics for 2010 show a growth of worldwide rail freight tonne-kilometres by 3.4% and worldwide rail passenger-kilometres by 3.5% from the previous year [1]. As for the longer-term outlook, strategic goals and predictions point to continued growth during the coming decades, as exemplified below.

In Europe, the European Commission in its 2011 transport white paper sets out strategic goals to transfer freight transportation from road: "30% of road freight over 300 km should shift to other modes such as rail or waterborne transport by 2030, and more than 50% by 2050" [2, 3].

In North America, the US Department of Transportation is forecasting an increase in tonnage transported by rail in the USA from 1890 megatonnes in 2010 to 2594 megatonnes in 2040, i.e. an increase of some 37% over 30 years [4].

To meet continued growth in transportation demand by rail, system capacity utilization will need to increase or the capacity be expanded. Increased capacity utilization generally reduces the margins against unplanned deviations in one part of the system spreading also to other parts, i.e. "knock-on" effects, thus risking a deterioration in service quality. Capacity expansion, on the other hand, may conflict with obstacles in the forms of budgetary constraints, land use and otherwise. Thus, in the face of rising transportation demand, it is important to use the railway system as efficiently as practicable, getting the highest productivity and quality with limited production assets, fixed as well as moving.

2 Literature review

What is capacity? What aspects of capacity are relevant to the railway as a system? How can one effectively raise capacity where needed?

Capacity in a general sense is the "capability to perform or produce" [5]. This is relevant, albeit imprecise. In the context of railway transportation, the capacity of a railway is often used to mean the number of trains that can travel over a given section of line during a given period of time.

The International Union of Railways' (UIC) Code 406 *Capacity*, which aims to establish a harmonised understanding of and to enable calculations of available railway infrastructure capacity, defines the capacity of any railway infrastructure as "the total number of possible paths in a defined time window, considering the actual path mix or known developments respectively and the IM's own assumptions; in nodes, individual lines or part of the network; with market-oriented quality" [6]. It is worth noting that Code 406 is directed specifically toward railway infrastructure managers (IM), and that it defines capacity as the number of possible train paths under given assumptions of traffic mix and quality. Code 406 points out that there are interdependencies between the number of trains per time interval, the average train speed, the stability against secondary delays, and the heterogeneity of different running times for different train types. It further distinguishes between maximum theoretical capacity under ideal circumstances, which are to be defined, and recommended standard values including margins for infrastructure and rolling stock (un)reliability.

Other recent analyses of the railway capacity concept include those by Landex and Abril.

Landex *et al*. in close agreement with UIC Code 406, characterize the capacity utilization of railways as being divided into four core elements, which are all dependent of each other: the number of trains, the average speed, the heterogeneity of the operation, and its stability (punctuality). Landex *et al*. further differentiate between maximum capacity and fundamental capacity, where the latter includes allowances for deviations [7, 8].

Abril *et al.* similarly define theoretical capacity as "the number of trains that could run over a route, during a specific time interval, in a strictly perfect, mathematically generated environment, with the trains running permanently and ideally at minimum headway" [9]. From this definition of theoretical capacity, Abril *et al.* derive practical capacity, used capacity and available capacity.

Thus it can be seen in a number of European sources that railway capacity is commonly defined as the number of trains in a given period and as a function of various traffic parameters.

In North America, in a review by the Canadian Transport Commission of the state of the art in defining and measuring railway capacity, Khan defines capacity at the corporate level as "the ability of the carrier to supply as required the necessary services within acceptable service levels and costs so as to meet the present and projected demand" [10, p.41], i.e. a relationship between capacity, quality and cost. Definitions are proposed for four different levels of capacity: engineering capacity, physical ultimate capacity, practical capacity and economic capacity.

Considering the complexity of the railway system and the interdependencies between its subsystems, such as its lines, yards, terminals, equipment, repair and maintenance facilities and other resources (manpower, funds, materials), Khan recognizes the need to define and measure capacity not only at the railway system level but also for its sub-systems and components. For a railway line, many of the quoted definitions of practical capacity focus on the number of trains per day and traffic conditions, but Khan also points out that this is limited by such factors as terrain, train size, tonnage, and operating procedures, among others. One of the quoted definitions focuses succinctly on the tonnage hauled and on the interrelationship between capacity, traffic mix and cost: "Capacity, in general terms, can be stated as the number of gross tons that can be handled over a definitive segment of track by a definitive number of trains of a definitive mix of service and size within a given period of time (according to CNR and CP), and at given cost" [10, p. 51]. Proposed measures of line capacity include: trains per day, gross tonne-miles per mile of track per day, average speed and running time.

Similarly, several measures of yard capacity are also proposed, focusing on the number of wagons or trains handled and the throughput time of each.

Khan points out that train and line capacity may be increased by raising the average speed or the load per wagon and per train, limited by the trade-offs between the power-to-weight ratio and train acceleration as well as the average speed on grades.

Krueger of Canadian National (CN) describes capacity in general terms: "Capacity is a measure of the ability to move a specific amount of traffic over a defined rail line with a given set of resources under a specific service plan. This could mean anything from the number of tons moved, speed of trains, on-time-performance, available track maintenance time, service reliability, or maximum number of trains per day that the subdivision can handle" [11].

Seeing trackage as the railway's largest asset, a parametric model of track capacity was developed by the CN as a simple decision support tool to

understand theoretical, practical, used and available capacity and to help improve track utilization and service quality. The model helps balance the demands for existing traffic, new traffic and maintenance access. The model expresses theoretical capacity as the theoretical maximum number of trains per day that can be moved over a subdivision in homogeneous traffic (all trains being the same) and without disruptions, but defines practical, used and available capacity in relation to the actual traffic mix. Assessing the available capacity for increased traffic, Krueger notes that there are trade-offs between different measures: "increasing tons moved could mean longer heavier trains which are slower and lead to reduced train speeds for the line, while increased train speeds require shorter and lighter trains, but potentially less tons moved for the same crews and locomotives".

The Association of American Railroads' (AAR) *National rail freight infrastructure capacity and investment study* of 2007 [12], based on a forecast increase of rail transportation demand in the USA until 2035 by some 88%, measured in tonnes, set out to identify the level of infrastructure improvements and investments necessary.

For each major rail freight corridor in the network, the study assesses its practical line capacity. This is done initially as number of trains per day, based on the number of parallel tracks (1, 2, 3, 4), the type of control system (TWC, ABS, CTC), and the degree of train type heterogeneity operating on the line. The average number of wagons per train is defined for each of the dominating train service types: intermodal, bulk, general merchandise and automobiles. For each commodity, the nominal capacity in tonnes per wagon is defined for the two prevailing wagon mass categories (119 tonnes vs. 130 tonnes gross rail load). Adjustments are made for typical proportions of empty movements per wagon type and for seasonal and day-to-day traffic variations. Thus, the physical infrastructure characteristics, the train size and the traffic patterns together give the calculated practical capacity in tonnes per year for each line, which is then compared to the forecast transportation demand. This ratio, the utilized capacity as a fraction of the practical capacity, is taken as a level of service grade, indicating the robustness to and ability to recover from disturbances. The AAR study also points to the importance of improving productivity by hauling more wagons per train and loading wagons closer to their capacity, to reduce the corridor expansion needs.

It is seen in the North American sources reviewed here, that while capacity is initially expressed as the number of trains that can travel over a section of line in a given period and under specific traffic conditions, this is also frequently converted to the tonnage that can be hauled. This is in contrast to the European sources that were reviewed, which analyze the number of trains extensively, but without giving much attention to the size of trains i.e. the loading capacity of each train, at least not in the same context.

3   Stakeholders' needs

The primary mission of the railway system, i.e. its reason for being, is to transport passengers and freight, in some cases in combination with other logistic services.

In analyzing the railway as a system, striving to improve its performance, the related actors and stakeholders and their needs must be considered. The ultimate system users are the freight shipper and consignee, whose needs have been surveyed extensively by Lundberg [13], and the passenger, but the needs of the train operator and infrastructure manager are equally important in being able to provide a good and viable service to the end users.

Surveying the needs of shippers and consignees, cost was identified by Lundberg as their highest priority. However, other aspects must also be fulfilled for the service to be attractive and competitive. Shippers generally find service reliability more important than a short transit time, although step effects such as the ability to provide overnight delivery may be important for specific origin-destination pairs.

Whether stakeholders belong to an integrated or to separate organizations, their needs depend on the functions they perform, e.g. passenger, freight shipper, freight consignee, train operator, infrastructure manager, maintainer, neighbour etc. The main stakeholders and their typical needs are shown in Table 1.

Table 1:   Railway stakeholders and their typical needs.

Stakeholder	Typical needs
Freight shipper	Cost, on-time supply of empty and pick up of loaded wagon, wagon in good order
Freight consignee	Cost, absence of damage, service reliability, on-time delivery
Passenger	Travel time, service frequency, cost, service reliability, comfort, service
Train operator	Cost, service reliability, equipment utilization, capacity, maintenance access
Infrastructure manager	Cost, service reliability, infrastructure utilization, capacity, maintenance access

Freight train operating costs for different scenarios have been broken down by Flodén [14] and Boysen [15], showing that a large portion of the costs is either independent of train size ("fixed") or increases less than proportionally to train size.

4   Capacity model

The capacity of the railway as a (complex) system should be analyzed in relation to the railway's primary mission, which is the transportation of freight and passengers, to the needs of its stakeholders, and to the applicable constraints.

The amounts transported can be characterized as tonnes, cubic metres and passengers. The production can be measured as tonne-kilometres, cubic-metre-kilometres and passenger-kilometres, respectively.
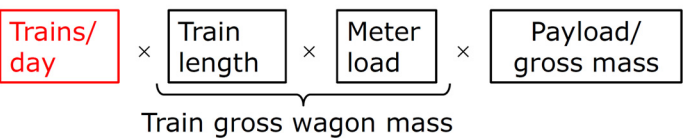
The capacity, then, is the maximum amount that can be produced in relation to the limiting constraints. Constraints include scarce production assets (e.g. infrastructure, rolling stock, staff) as well as consumables used up (e.g. energy, time). Low cost is a high priority for many stakeholders, and this, in a capital-intensive industry, requires high utilization of the available production assets.

The train as the production unit is decisive to the entire railway system: With commodities being transported by trains, the load capacity per train is one factor determining the railway's overall system capacity. The loading capacity per train will be limited either by its useful volume, in the case of low-density commodities, or by its load (mass) limit, in the case of high-density commodities, i.e. the load will either "cube out" or "weigh out" depending on which limit is reached first. Both cases must be considered, so a model for each case is needed.

To analyze the capacity per train, we depend primarily on parameters that are limited at the train level, such as maximum train length (metres), metre load (tonnes/metre) and useful cross section (square metres). (This is in contrast to wagon design, which depends primarily on parameters limited at the individual wagon level, such as number of axles, axle load, wagon length and, again, useful cross section.) The models for mass transport capacity and volume transport capacity, both per train, are shown in both of Figures 1 and 2.

Now that the loading capacity per train is expressed as tonnes and cubic metres in both of Figures 1 and 2, we still need measures of production that are related to the system mission, i.e. transportation in tonne-kilometres and cubic-
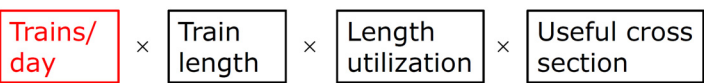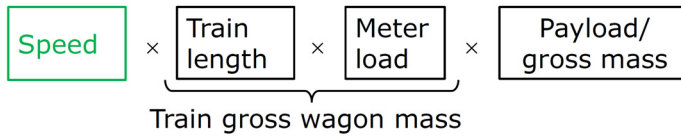

Figure 1:   General model of railway system transport capacity – infrastructure (fixed asset) perspective.

- Mass transport capacity (high density goods)

Train gross wagon mass

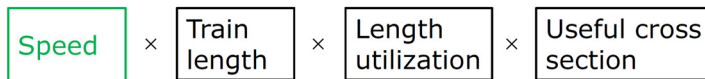- Volume transport capacity (low density goods)

Figure 2:   General model of railway system transport capacity – train operating (moving asset) perspective.

metre-kilometres, and measures of productivity that are related to production rate or production asset utilization.

Assets can be considered as fixed assets or moving assets; fixed production assets (or infrastructure) include the line with its track, electrification, signalling and communication systems (as applicable), yards and terminals, whereas moving assets include the locomotive(s), wagons, on-board train crew and lading.

From the perspective of the infrastructure manager and the fixed assets, the production rate per line section, the amount of load that can pass a given point during a given period of time, is the number of trains that can pass times the load capacity per train.   Transport capacity = transportable amount/unit time = number of trains/unit time × load per train. This is shown in Figure 1.

From the perspective of the train operator and the moving assets, the production rate per train, the amount of load that can be hauled a given distance during a given period of time, is the (average) train speed times the load capacity per train. Transport capacity = transportable amount × distance/unit time = train speed × load per train. This is shown in Figure 2.

Thus, the four capacity models shown in Figures 1 and 2 model capacity as being proportional to train frequency and speed, respectively, and to load capacity per train ("train size").

However, there are several trade-offs between train frequency (or inversely, headway) and speed and train size, and this is why capacity should be analyzed in one context, preferably including all of the above.

Similar models can be constructed for passenger service, expressing the number of passenger slots (seats or standees) per train in constrained design parameters such as number of seats across, longitudinal seat pitch, length of passenger saloon, number of floors per coach, number of coaches per train etc.

5   Trade-offs between capacity parameters

In many cases, however, there are trade-offs between capacity parameters, i.e. if one is increased, others will decrease. To maximize the product of load per train and train frequency or speed, therefore, these factors must be considered jointly and not separately.

Trade-offs between train speed, train frequency and the amount of load per train ("train size") include:

1) locomotive power: trade-off between train mass and speed for given line gradient;
2) locomotive tractive effort: lower acceleration for higher train mass;
3) wagon braking performance: lower wagon load limit for higher speed;
4) train braking performance: shorter train length for higher speed;
5) train braking performance: longer stopping distance and longer headway for higher speed.

In a homogeneous traffic situation, all trains running at the same speed and not catching up with each other, it is the product of load per train and train speed or train frequency that should be maximized to achieve maximum system capacity.

In contrast, in a heterogeneous traffic situation, trains with different speeds or different stopping patterns may interfere with each other, and the degree of heterogeneity is one of the factors affecting capacity; a high degree of heterogeneity necessitates longer headway, thus reducing the number of trains that can run in a given period and requiring passing (overtaking).

Thus, where traffic is heterogeneous, minimizing the heterogeneity e.g. reducing speed differences will increase capacity.

6   Populating and filtering the model

The model helps visualize what parameters are important. Populating it with data for the constituent factors makes it easy to get an approximate grip of the range of improvements feasible. In Figure 3, the model is populated with parameter data, of which blue figures represent the range presently in use in Sweden.

Now that an overview has been gained, it is easy to filter the model to focus attention on those factors giving the highest effect. An example of this is shown in Figure 5, which lists only those factors which individually would increase capacity by some 30% or more. By combining several factors, higher overall effect can be achieved.
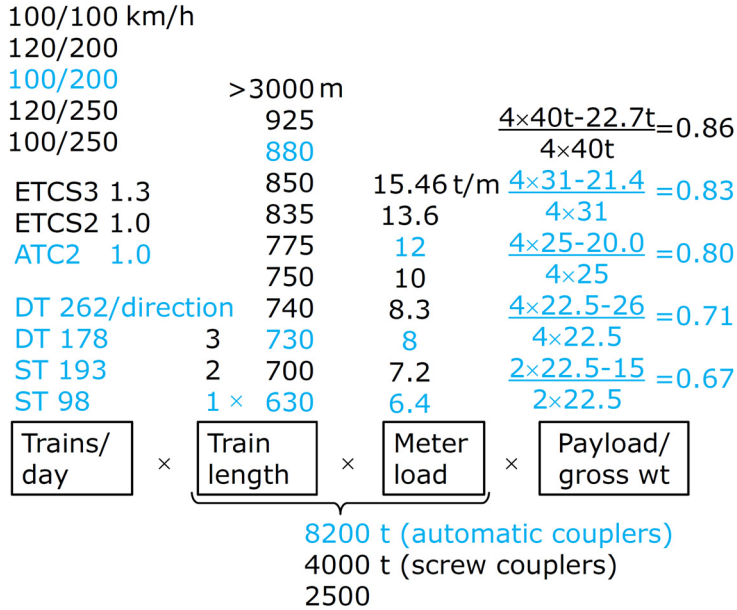
100/100 km/h
120/200
100/200
120/250
100/250

ETCS3 1.3
ETCS2 1.0
ATC2 1.0

DT 262/direction
DT 178
ST 193
ST 98

>3000 m
925
880
850
835
775
750
740
3   730
2   700
1 × 630

15.46 t/m
13.6
12
10
8.3
8
7.2
6.4

$$\frac{4\times40t-22.7t}{4\times40t}=0.86$$

$$\frac{4\times31-21.4}{4\times31}=0.83$$

$$\frac{4\times25-20.0}{4\times25}=0.80$$

$$\frac{4\times22.5-26}{4\times22.5}=0.71$$

$$\frac{2\times22.5-15}{2\times22.5}=0.67$$

Trains/ day	×	Train length	×	Meter load	×	Payload/ gross wt

8200 t (automatic couplers)
4000 t (screw couplers)
2500

Figure 3:   Populated model of railway system mass transport capacity – infrastructure perspective (DT=double track, ST=single track, blue=Sweden).

IN-DFC 17 m²
US-DS 15.0
AAR K 13.5
3.60×4.90 13.3
SE-C 13.1
3.40×4.90 12.6
AAR F 12.3
3.15×4.90 11.7
FI-KU 10.8
UIC GC 10.0
P 450 9.1
SE-A 8.3
P 400 7.8
DE-G2 7.3
DE-G1 6.7

100/100 km/h
120/200
100/200
120/250
100/250

ETCS3 1.3
ETCS2 1.0
ATC2 1.0

DT 262/direction 740   22m/23.26m=0.95
DT 178   3   730   13.88/15.14=0.92
ST 193   2   700   27.2/34.03=0.80
ST 98   1 × 630   13.6/18.34=0.74

>3000 m
925
880
850
835
775
750

Trains/ day	×	Train length	×	Length utilization	×	Useful cross section*

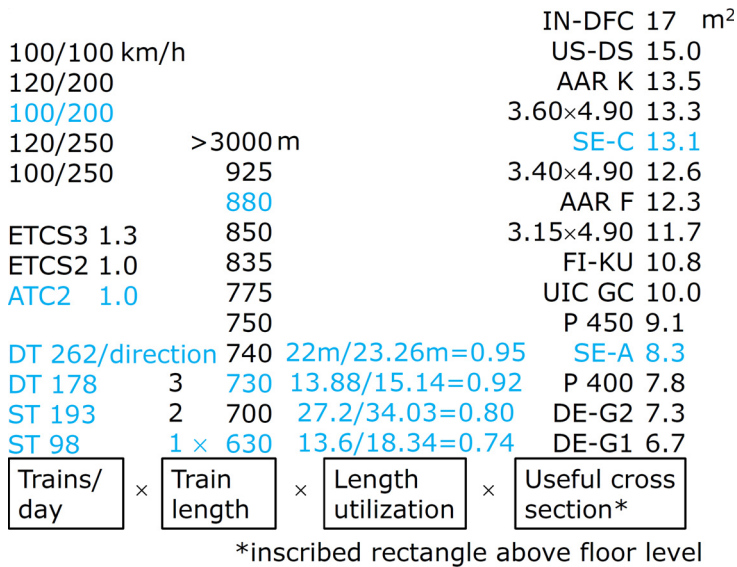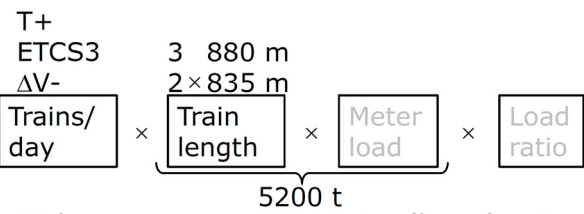*inscribed rectangle above floor level

Figure 4:   Populated model of railway system volume transport capacity – infrastructure perspective (DT=double track, ST=single track, blue=Sweden).

- Mass transport capacity (high density goods)

T+
ETCS3   3 880 m
ΔV-   2×835 m

| Trains/ day | × | Train length | × | Meter load | × | Load ratio |

5200 t

- Volume transport capacity (low density goods)

T+
ETCS3   3 880 m
ΔV-   2×835 m   Flat top

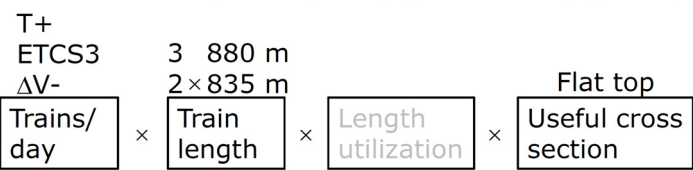| Trains/ day | × | Train length | × | Length utilization | × | Useful cross section |

Figure 5:   Filtered model of railway system volume transport capacity – infrastructure perspective, showing only factors that each would boost capacity by 30% or more over present common levels (in Sweden) (T=number of tracks, ΔV=speed difference).

7   Case studies

The model has been applied to several case studies, and has demonstrated results that were not intuitively obvious.

7.1 Case study #1: capacity effects of higher axle load

In the first case study, the effects on the transportation capacity of iron ore were investigated for an axle load increase from 25 tonnes to 30 tonnes. The loading capacity per wagon is determined by the number of axles, axle load and wagon tare mass, whereas the loading capacity per train is determined mainly by the train length and metre load. In this case, the new, higher-axle load wagons were made longer, the wagon metre load was actually reduced slightly, and the permissible metre load for the track remained unchanged.
    Referring to the net payload, it was shown that the mass loading capacity (tonnes) per wagon increased by some 23%, i.e. slightly higher than the axle load increase of 20%. The mass loading capacity (tonnes) per metre of train length increased by 0.5%, i.e. nearly unchanged, which is because the metre load rating of the track at 12 tonnes/metre was already fully used, and was not raised. Higher loading capacity (tonnes) per train was also achieved, but this is the result of increased train length and could have been achieved without raising the axle load.
    This case is interesting because savings materialized in the area of rolling stock, from being able to carry a given tonnage with fewer wagons, whereas costs were incurred in the area of track upgrading. This can present particular challenges where the wagon owner, train operator and infrastructure manager are

different organizations: How can the cost impact and savings be shared equitably, and how should the system be designed for overall optimum efficiency?

7.2 Case study #2: capacity effects of faster freight trains

In the second case study, the effects on the transportation capacity of modern intermodal trains were investigated for a possible speed increase from 100 km/h to 120 km/h. Here, the loading capacity and transport capacity per train are influenced by trade-offs between train speed and train length, as well as between train speed and the load limit (tonnes) per wagon, due to limitations in braking performance (stopping distance and wheel thermal capacity).

It was shown that the volume loading capacity (cubic metres) per train would decrease by some 29% and the volume transport capacity (cubic-metre-km/h) decrease by some 15% at the higher speed. Further, the mass loading capacity (tonnes) per train would decrease by some 42% and the mass transport capacity (tonne-km/h) per train decrease by some 31%.

Whether such a speed increase would still be desirable, despite the reduced loading and transport capacity per train, would also depend on other specific factors, such as the traffic mix and any speed difference from other trains, which may show variation e.g. between day and night, between weekdays and weekends et cetera.

8 Discussion

The demonstrated model is easy to use. Relative comparisons of the effects on transportation capacity from different candidate measures can be made easily, with data accounting for the differences. Absolute comparisons of the attainable transportation capacity from specific measures against absolute capacity targets require more complete data.

The model can quickly give an approximate assessment, including trade-off effects, and can point to measures that may be worth investigating more in detail by other methods, e.g. by traffic simulation.

The two case studies related above gave mixed results on important capacity indicators: positive, negligible, or negative. The magnitudes of some of the effects are significant, up to a 42% decrease in one case.

These results may in part be contrary to expectation, and would not have been demonstrated without systematic quantitative application of the trade-offs at several levels.

9 Conclusions

In railway capacity planning, when comparing the efficacy of alternative investment or operating scenarios with each other or with a particular demand forecast, it is not enough to assess capacity merely in terms of number of trains per hour (or day). Rather, one should revert to the definition of capacity, what is

able to be transported, which in freight transportation is cubic volume and mass and in passenger transportation number of seats and standing area. Going back to these definitions is especially important when the alternatives to be assessed have mixed effects and the combined effect is not obvious, e.g. when an increase in the speed and number of trains is also associated with a reduction in train size.

A general model of railway transportation capacity has been developed, which models (i) mass transportation capacity and (ii) volume transportation capacity, expressed as tonnes or tonne-km and cubic meters or cubic-metre-km, respectively, per unit of time (hour or day). Variants of the model address the productivity of the scarce resource of infrastructure or train operator, respectively, i.e. the railway link or the train.

The model takes into account physical design parameters such as dimensions and load ratings, and utilization factors such as length utilization and payload as a fraction of gross load.

Graphically, in a block diagram, the model visualizes the constituent parameters, their relationship and the range of available magnitudes for each parameter. By visualizing this in a compact form the model helps focus attention on what parameters are more significant, as well as the overall capacity effect attainable by combining parameters. Thus the model offers a simple and quick way of assessing which of several alternatives is more effective or whether a proposed improvement or combination of improvements is potentially able to fulfil a forecast capacity need.

Application of the model to principally important cases has given results that are not intuitively obvious, and which may not have been reached otherwise:

1) Higher axle load results in fewer wagons being needed to carry a given payload tonnage, but does not necessarily contribute to higher transportation capacity per train or for the railway system as a whole.
2) Higher freight train speeds can result in significantly lower loading capacity and transport capacity per train, and in a homogeneous traffic situation even for the railway system as a whole.

Decisions on railway infrastructure and rolling stock investment as well as on railway operating strategies should be tested against the general transportation capacity model to assess their potential impacts on transportation capacity.

References

[1]   UIC World Rail Statistics. International Union of Railways: Paris, 2011 July 27. http://www.uic.org/IMG/pdg/cp18_uic_stats_2010_en-2.pdf

[2]   *Roadmap to a Single European Transport Area – Towards a Competitive and Resource Efficient Transport System.* European Commission: Brussels, p. 9, 2011 March 28.

[3]   Nelldal B.L. and Andersson E., Mode shift as a measure to reduce greenhouse gas emissions. *Transport Research Arena*: Athens, 2012 April 23–26.

[4] Freight Analysis Framework, version 3.2. Center for Transportation Analysis, Oak Ridge National Laboratory: Knoxville, 2011 December 1, updated 2012 June 18; cited 2012 June 18. http://faf.ornl.gov /fafweb/Extraction1.aspx

[5] Webster's Online Dictionary. Cited 2012 July 9. http://www.websters-online-dictionary.org

[6] *UIC Code 406: Capacity*, 1st issue: Paris, International Union of Railways: Paris, p. 6, 2004 June.

[7] Landex A., Kaas A.H., Schittenhelm B. and Schneider-Tilli J., Evaluation of railway capacity. *Traffic Days at Aalborg University*: Aalborg, pp. 3–8, 2006 August 28–29.

[8] Landex A., Capacity statement for railways. *Traffic Days at Aalborg University*: Aalborg, pp. 9–13, 2007 August 27–28.

[9] Abril M., Barber F., Ingolotti L., Salido M.A., Tormos P. and Lova A., An assessment of railway capacity. *Transportation Research Part E*, 44, p. 776, 2008.

[10] Khan, A.M., *Railway Capacity Analysis and Related Methodology*, Canadian Transport Commission: Ottawa, 1979.

[11] Krueger H., Parametric modeling in rail capacity planning. *Winter Simulation Conference*: Phoenix, p. 1195, 1999.

[12] Grenzeback L.R., Hunt D.T. and Beagan D.F., *National Rail Freight Infrastructure Capacity and Investment Study*, Cambridge Systematics, Association of American Railroads: Cambridge, 2007 September.

[13] Lundberg S., Freight customers' valuations of factors of importance in the transportation market, Royal Institute of Technology: Stockholm, 2006.

[14] Flodén J., Rail freight costs: Some basic cost estimates for intermodal transport, University of Gothenburg: Gothenburg, 2011 March.

[15] Boysen H.E., More efficient freight transportation through longer trains. *Transport Forum*: Linköping, 2012 January 11-12. http://www.slideshare.net/hansboysen/session-42-hans-boysen and http://www.slideshare.net/hansboysen/session-42-hans-boysen-11138646